

Tabla de Contenido

1. Introducción	1
1.1. Descripción de mercado del USDCLP	1
1.2. Descripción del problema	2
1.3. Justificación del trabajo de tesis	4
1.4. Resultados esperados	5
1.5. Objetivo principal y específicos	6
1.6. Sustento de la metodología de aprendizaje por refuerzo	7
1.7. Elementos complementarios	7
1.8. Plan de trabajo	8
1.9. Trabajos relacionados	8
2. Marco teórico	10
2.1. Recapitulación	10
2.2. Situación base y aspectos de mercado	12
2.3. Solución desarrollada	13
2.4. Integración de los elementos esenciales de la solución	14
2.5. Totalidad de un tipo de operación a ser modelado	16
2.6. Supuesto de simetría en las operaciones de venta y compra	16
3. Concepción de la solución	17
3.1. Requisitos y restricciones a la solución	17
3.2. Perfiles de usuario	18

3.3. Introducción al aprendizaje por refuerzo	18
3.3.1. Definición de estados	18
3.3.2. Estado inicial, iteración de estados y estado terminal	19
3.3.3. Acciones	20
3.3.4. Política de decisión	20
3.3.5. Recompensa	20
3.3.6. Exploración e intensificación	21
3.3.7. Rescatando lo aprendido	21
3.3.8. Aprendizaje cíclico	21
3.4. Interacción entre los distintos componentes	22
3.5. Valorización de estados y su relación a las acciones	24
3.6. Política de decisión y su relación con las metodologías de aprendizaje estructurado	25
3.7. Relación entre la valorización de estados y de acciones - estados	27
3.8. Optimización de la política de decisión	28
3.9. Origen de la metodología de resolución	29
3.10. Métodos de aprendizaje por refuerzo	30
3.11. Características método diferencia temporal	31
3.12. Justificación implementación metodología diferencia temporal	32
3.13. Definición metodología de diferencia temporal	32
3.14. Aplicación de metodología de diferencia temporal: Q-Learning - SARSA	33
3.15. Valorización de producto financieros mediante procesos estocásticos	34
3.16. Generador de precios estocásticos de tipo de cambio	36
3.17. Abstracción del impacto de volumen de una transacción	38
3.18. Métricas de riesgo de mercado	39
3.19. Definición de la métrica de Valor en Riesgo	40
3.20. Utilización del VaR en la metodología de Aprendizaje por Refuerzo	40

4. Implementación de la solución	42
4.1. Ejemplificación de una trayectoria del problema	42
4.1.1. Decisiones	42
4.1.2. Estados y estado inicial	42
4.1.3. Recompensa	44
4.1.4. Transición de estados	45
4.2. Aplicando métricas de precio y recompensa acordes con la realidad	46
4.3. Implementación de la solución	47
4.3.1. Identificación y calibración de los parámetros del modelo	47
4.4. Desarrollo del algoritmo	48
4.4.1. Activo Financiero - USDCLP	48
4.4.2. Agente - Parámetros	51
4.4.3. Agente - Decisión	53
4.4.4. Agente - Innovación	54
4.4.5. Agente - Aprendizaje	55
5. Prueba conceptual del modelo y metodología	57
5.1. Parámetros del Ejercicio	58
5.2. Ejecución del algoritmo	59
5.3. Resultados del Ejercicio	61
5.4. Interpretación de los resultados	64
5.5. Heurística como metodología comparativa	65
6. Recapitulación y conclusiones	67
6.1. Conclusiones complementarias del presente trabajo	68
6.1.1. Materialización de aprendizaje	68
6.1.2. Discriminación en riesgo-retorno	69
6.1.3. Flexibilidad, modularidad y adaptabilidad de la solución	69

6.1.4. Impacto de la aleatoriedad en el aprendizaje	69
6.1.5. Capacidad para analizar acotados dominios del problema	69
6.1.6. Estados con limitada experiencia y extracción de aprendizaje	70
6.1.7. Comparativa frente a situación basal	70
6.1.8. Capacidad para modelar eventos probabilísticos	70
6.1.9. Algoritmo de caja negra	70
6.1.10. Reducido número de variables explicativas	70
6.1.11. Rapidez de cálculo de resultados	71
6.1.12. Complejidad de modelar precios intradías	71
6.2. Trabajo futuro	72
Glosario	73
Bibliografía	75